Human deadenylating nuclease, its preparation and use

Description

The present invention relates to a human deadenylating nuclease (DAN), to its coding nucleic acid and to its preparation and use.

The intracellular concentration of mRNA appears to be controlled by its rate of degradation as well as by its rate of production. In this connection, mRNAs appear not to be degraded by a random nucleolytic event but, rather, by specific mechanisms and by rates of degradation which are specific for the given RNA. Currently, two different degradation pathways are known in which the poly(A) tails play a particular role. In E. coli, poly(A) tails are appended to mRNAs or mRNA fragments which are generated by RNAse E. The poly(A) tail appears to function as a relatively unstructured unit which is responsible for the attachment of a complex which, in addition to other proteins, contains a progressive 3'-exonuclease, which is a polynucleotide phosphorylase, and an RNA-dependent ATPase which helps the exonuclease to circumvent inhibitory secondary structures in the mRNA. A similar mechanism appears to operate in chloroplasts.

In eukaryotes, exonucleolytic truncation of the poly(A) tail likewise initiates degradation of many but not all mRNAs. In contrast to the process in bacteria, the poly(A)-degrading exonuclease does not appear, in carrying out its degradation, to continue into the 3'-UTR and the coding sequence. Instead, after having degraded the poly(A) tail, the eukaryotic exonuclease appears to stop. In S. cerevisiae, the second step of mRNA degradation involves removal of the 5' cap structure by a specific pyrophosphatase. However, the CAP is only removed after the the poly(A) tail has been shortened down to approx. 10-15 nucleotides. Removal of the CAP structure makes the mRNA accessible to 5'-exonucleases, whose most important representative is encoded by the XRN1 gene.

While it is relatively easy to observe the initial deadenylation process in mammalian cells, investigation of the additional steps is rendered more difficult by the rapidity of the subsequent degradation processes and the absence of analyzable intermediates. However, indirect evidence suggests that the CAP structure is removed in the second step. A homolog of the

XRN1 gene has been described in the mouse. In general, stable mRNAs are deadenylated slowly and unstable mRNAs are deadenylated rapidly. Rapid degradation depends on the presence of particular destabilizing sequences in the 3'-UTR or in the coding sequence. Furthermore, mutations in these sequences lead not only to the mRNA being stabilized but also to deadenylation being slowed down. By contrast, inactivation of a stabilizing element in α-globulin mRNA leads to instability and to accelerated deadenylation. These data provide clear support for degradation of mRNA being controlled by the deadenylation rate.

The deadenylation of particular mRNAs can lead to inactivation of translation as well. This can be explained by the importance of the poly(A) tail for the initiation of translation. Deadenylation as a mechanism for regulating translation plays a crucial role in the maturation of oocytes and in early embryogenesis in many animal species. For example, in Drosophila, the polarity of the embryo is regulated by deadenylation of the so-called hunchback mRNA.

In vertebrates, three deadenylation reactions can in principle be distinguished in oocyte maturation and early embryogenesis:

1. In immature oocytes, most mRNAs are stored in a translationally inactive form in which they possess short poly(A) tails. An example of this is the mRNA for tPA (tissue plaminogen activator) in mice. This mRNA is initially given a long poly(A) tail in the normal polyadenylation reaction, with this tail then being truncated by deadenylation into an oligo(A) tail. This truncation is regulated by sequences in the 3'-UTR.
2. During maturation of oocytes, deadenylation is used to inactivate particular mRNAs which originally had long poly(A) tails and were active in early egg development. The deadenylation does not depend on specific sequences in the mRNA. All mRNAs are deadenylated unless they are protected by an active adenylation process.
3. During early embryogenesis, certain mRNAs are deadenylated in a specific reaction which likewise requires specific sequences in the 3'-UTR.

All three deadenylation reactions takes place in the cytoplasm. However, in contrast to the situation in somatic cells, the oligoadenylated or

deadenylated mRNAs remain stable in oocytes and are either adenylated once again or only degraded in subsequent development stages.

The enzymes which are responsible for these reactions have so far not been identified unambiguously. It has been shown in yeast that degradation of poly(A) depends directly or indirectly on the poly(A)-binding protein (PAB1). Although a PAB1-dependent poly(A)-nuclease (PAN) has been purified (Sachs, A.B. & Deardorff (1992) Cell, 70, 961; Lowell, J.E. et al. (1992) Genes Dev., 6, 2088), it has only been possible to identify minor defects in deadenylation in mutants of the genes concerned (PAN2 and PAN3) (Boeck, R. et. Al. (1996) 271 (1), 432; Brown, C. et al. (1996) 16 (10) 5744).

Körner, Ch. G. & Wahle, E. (1997) J. Biol. Chem., 272, 10448, No. 16 describe an Mg^{2+}-dependent poly(A)-specific 3'-exoribonuclease from calf thymus which has a molecular weight of 74 kDa and which has also been referred to as being a deadenylating nuclease (DAN). The calf thymus DAN which has been described is stimulated by the cytoplasmic poly(A)-binding protein I (PAB I) under defined reaction conditions (Körner, Chr. & Wahle, E. (1997), see above) and preferentially poly(A) as its substrate.

The object of the present invention was to make available a human poly(A)-specific 3'-exoribonuclease.

Surprisingly, a clone which encodes a human deadenylating nuclease (human DAN) in accordance with the present invention, but which had not been characterized in detail and which had only been sequenced terminally, has now been found in a gene library.

The present invention therefore relates to a nucleic acid which encodes a human DAN having an amino acid sequence as depicted in SEQ 11 or a functional variant thereof, and parts thereof having at least 8 nucleotides, preferably having at least 15 or 20 nucleotides, in particular having at least 100 nucleotides, very particularly having at least 300 nucleotides (subsequently termed "nucleic acid according to the invention").

The complete nucleic acid encodes a protein having 639 amino acids and a molecular mass of 73.5 kDa. Expression of the nucleic acid in E. coli gave rise to a protein which displays enzymic activities which are similar to

those displayed by the DAN described by Körner, Ch. G. & Wahle, E. (1997, see above). Other experiments performed in accordance with the present invention confirmed that the nucleic acid is a nucleic acid which encodes a human DAN. The nucleic acid according to the invention was deposited in the DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen [German collection of microorganisms and cell cultures] GmbH, Mascheroder Weg 1b, 38124 Brunswick under number DSM 12075 on March 25, 1998.

In a preferred embodiment, the nucleic acid according to the invention is a DNA or RNA, preferably a double-stranded DNA, and in particular a DNA having a nucleic acid sequence as depicted in SEQ 12 from pos. 58 to pos. 1977. The two positions determine, according to the present invention, the beginning and the end of the coding region.

According to the present invention, the term "functional variant" is understood as meaning a nucleic acid which is functionally related to the human DAN-encoding nucleic acid and is, in particular, of human origin. Examples of related nucleic acids are nucleic acids from different human cells or tissues, or allelic variants. The present invention likewise encompasses variants of nucleic acids which can be derived from different human individuals.

In the broader sense, the term "variants" is understood, according to the present invention, as meaning nucleic acids which exhibit a homology, in particular a sequence identity, of approx. 60%, preferably of approx. 75%, in particular of approx. 90% and, very particularly, of approx. 95%.

The parts of the nucleic acid according to the invention can, for example, be used for preparing individual epitopes, as probes for identifying other functional variants or as antisense nucleic acids. For example, a nucleic acid composed of at least approx. 8 nucleotides is suitable for use as an antisense nucleic acid, a nucleic acid composed of at least approx. 15 nucleotides is suitable for use as a primer in the PCR method, a nucleic acid composed of at least approx. 20 nucleotides is suitable for identifying further variants, and a nucleic acid composed of at least approx. 100 nucleotides is suitable for use as a probe.

In another preferred embodiment, the nucleic acid according to the invention contains one or more noncoding sequences and/or a poly(A) sequence. The noncoding sequences are, for example, intron sequences or regulatory sequences, such as promoter or enhancer sequences, for regulated expression of the human DAN-encoding gene.

In another embodiment, the nucleic acid according to the invention is therefore contained in a vector, preferably in an expression vector or a vector which is effective in gene therapy.

The expression vectors can, for example, be prokaryotic or eukaryotic expression vectors. Examples of prokaryotic expression vectors for expression in E. coli are the T7 expression vector pGM10 (Martin, 1996), which encodes an N-terminal Met-Ala-His6 tag, which enables the expressed protein to be advantageously purified by way of an Ni^{2+}-NTA column. Examples of eukaryotic expression vectors which are suitable for expression in Saccharomyces cerevisiae are the vectors p426Met25 and p426GAL1 (Mumberg et al. (1994) Nucl. Acids Res., 22, 5767), while examples of such vectors which are suitable for expression in insect cells are baculovirus vectors as disclosed in EP-B1-0127839 or EP-B1-0549721, and for expression in mammalian cells are SV40 vectors, with these vectors being generally available.

In general, the expression vectors also contain regulatory sequences which are suitable for the host cell, such as the trp promoter for expression in E. coli (see, e.g., EP-B1-0154133), the ADH-2 promoter for expression in yeasts (Russel et al., (1983), J. Biol. Chem. 258, 2674), the baculovirus polyhedrin promoter for expression in insect cells (see, e.g., EP-B1-0127839) or the early SV40 promoter or LTR promoters, e.g. from MMTV (mouse mammary tumor virus; Lee et al. (1981) Nature, 214, 228).

Examples of vectors which are effective in gene therapy are virus vectors, preferably adenovirus vectors, in particular replication-deficient adenovirus vectors or adenoassociated virus vectors, e.g. an adenoassociated virus vector which is composed solely of two inserted terminal repeat (ITR) sequences.

Examples of suitable adenovirus vectors are described in McGrory, W.J. et al. (1998) Virol. 163, 614; Gluzman, Y. et al. (1982) in "Eukaryotic Viral

Vectors" (Gluzman, Y. ed.) 187, Cold Spring Harbor Press, Cold Spring Harbor, New York; Chroboczek, J. et al. (1992) Virol. 186, 280; Karlsson, S. et al. (1986) EMBO J., 5, 2377 or WO 95/00655.

Suitable adenoassociated virus vectors are described, for example, in Muzyczka, N. (1992) Curr. Top. Microbiol. Immunol. 158, 97; WO 95/23867; Samulski, R.J. (1989) J. Virol. 63, 3822; WO 95/23867; Chiorini, J.A. et al. (1995) Human Gene Therapy 6, 1531 or Kotin, R.M. (1994) Human Gene Therapy 5, 793.

Vectors which are effective in gene therapy can also be obtained by complexing the nucleic acid according to the invention with liposomes. Lipid mixtures as described in Felgner, P.L. et al. (1987) Proc. Natl. Acad. Sci, USA 84, 7413; Behr, J.P. et al. (1989) Proc. Natl. Acad. Sci. USA 86, 6982; Felgner, J.H. et al. (1994) J. Biol. Chem. 269, 2550 or Gao, X. & Huang, L. (1991) Biochim. Biophys. Acta 1189, 195 are suitable for this purpose. When the liposomes are being prepared, the DNA is bonded ionically to the surface of the liposomes in a ratio which is such that a positive net charge remains and the DNA is completely complexed by the liposomes.

The nucleic acid according to the invention can, for example, be synthesized chemically, e.g. by the phosphotriester method (see, e.g., Uhlman, E. & Peyman, A. (1990) Chemical Reviews, 90, 543, No. 4), either on the basis of the sequence disclosed in SEQ 12 or on the basis of the peptide sequence disclosed in SEQ 11 and making use of the genetic code. Another possibility of obtaining the nucleic acid according to the invention is that of using a suitable probe (see, e.g., Sambrook, J. et al. (1989) Molecular Cloning. A laboratory manual. 2nd Edition, Cold Spring Harbor, New York) to isolate it from a suitable gene library, for example from a human gene library. Examples of suitable probes are single-stranded DNA fragments which are of a length of from approx. 100 to 1000 nucleotides, preferably of a length of from approx. 200 to 500 nucleotides, in particular of a length of from approx. 300 to 400 nucleotides, and whose sequence can be derived from the nucleic acid sequence depicted in SEQ 12.

The present invention furthermore also relates to the polypeptide itself having an amino acid sequence as depicted in SEQ 11 or a functional

variant thereof, and parts thereof having at least six amino acids, preferably having at least 12 amino acids, in particular having at least 65 amino acids and very particularly having 638 amino acids (subsequently termed "polypeptide according to the invention"). For example, a polypeptide which is approx. 6-12 amino acids in length, preferably approx. 8 amino acids in length, can contain an epitope which, after having been coupled to a carrier, is used for preparing specific polyclonal or monoclonal antibodies (in this regard, see, e.g., US 5,656,435). Polypeptides whose length is at least approx. 65 amino acids can also be used directly, without any carrier, for preparing polyclonal or monoclonal antibodies.

Within the meaning of the present invention, the term "functional variant" is understood as denoting polypeptides which are functionally related to the peptide according to the invention, i.e. which exhibit a poly(A)-specific 3'-exoribonuclease activity and which are preferably active under two different reaction conditions. In the first reaction condition, the activity is completely dependent on the presence of spermidene when salt is absent. By contrast, the activity of the enzyme is dependent on the salt concentration when spermidene is absent. Furthermore, it is possible to observe stepwise degradation of the poly(A) tail under defined reaction conditions when PAB1 is present (see also Körner, Chr. G. & Wahle, E. (1997), see above). In particular, the lengths of the predominant degradation products differ by approx. 30 nucleotides. Variants are also understood as meaning allelic variants or polypeptides which are derived from other human cells or tissues. They are also understood as meaning polypeptides which are derived from different human individuals.

In the broader sense, they are also understood as meaning polypeptides which possess a sequence homology, in particular a sequence identity, with the polypeptide having the amino acid sequence depicted in SEQ 11 of approx. 70%, preferably of approx. 80%, in particular of approx. 90%, very particularly of approx. 95%. They furthermore also include deletion of the polypeptide in the range of approx. 1 - 60, preferably of approx. 1 - 30, in particular of approx. 1 - 15, very particularly of approx. 1 - 5 amino acids. For example, the first amino acid, i.e. methionine, can be missing without the function of the polypeptide being significantly altered. In addition, they also include fusion proteins which contain the above-described polypeptides according to the invention, with the fusion proteins already themselves possessing the function of a human DAN or only being able to

acquire the specific function after the fusion moiety has been eliminated. Very particularly, they include fusion proteins having a content of, in particular, nonhuman sequences of approx. 1 - 200, preferably approx. 1 - 150, in particular of approx. 1 - 100, very particularly of approx. 1 - 50 amino acids. Examples of nonhuman peptide sequences are prokaryotic peptide sequences, for example from E. coli galactosidase or a so-called histidine tag, e.g. a Met-Ala-His_6 tag. A fusion protein containing a so-called histidine tag is particularly advantageously suitable for purifying the expressed protein by way of metal ion-containing columns, for example by way of an Ni^{2+}-NTA column. "NTA" stands for the chelating agent nitrilotriacetic acid (Qiagen GmbH, Hilden).

The parts of the polypeptide according to the invention represent, for example, epitopes which can be specifically recognized by antibodies.

By comparing with known nucleases, it was found that the polypeptide according to the invention is a member of the RNaseD family. Figure 4 shows the conserved amino acids of the Exo I, Exo II and Exo III motifs which are characteristic of this class of exonucleases. Other conserved amino acids have been given a gray background. The three acidic amino acid side chains, i.e. two in the Exo I domain and one in the Exo III domain, are involved directly in binding the two Mg^{2+} ions which participate in the enzymic hydrolysis. A third acidic amino acid side chain, which is located in the Exo II domain, binds one of the metal ions by way of hydrogen bond molecules. All these amino acid residues are highly conserved within the family and are also found in the human DAN according to the invention. The polypeptide according to the invention can therefore be described as being a poly(A)-specific 3'-exoribonuclease belonging to the RNaseD family.

The polypeptide according to the invention is prepared, for example, by expressing the nucleic acid according to the invention in a suitable expression system, as already described above, using methods which are well known to the skilled person. Examples of suitable host cells are the E. coli strains DH5, HB101 and BL21, the yeast strain Saccharomyces cerevisiae, the insect cell line Lepidopteran, e.g. from Spodoptera Frugiperda, or animal cells, such as COS, Vero, 293 and HeLa cells, all of which are available generally.

In particular, said parts of the polypeptide can also be synthesized by classical peptide synthesis (Merrifield technique). They are particularly suitable for obtaining antisera which can be used for screening suitable gene expression libraries for the purpose of gaining access to other functional variants of the polypeptide according to the invention.

The present invention therefore also relates to a process for preparing a polypeptide according to the invention, with a nucleic acid according to the invention being expressed in a suitable host cell and being isolated, where appropriate.

The present invention furthermore also relates to antibodies which react specifically with the polypeptide according to the invention, with the abovementioned parts of the polypeptide either themselves being immunogenic or being able to be made immunogenic, or to have their immunogenicity increased, by being coupled to suitable carriers, for example bovine serum albumin.

The antibodies are either polyclonal or monoclonal. Their preparation, which is also part of the subject-matter of the present invention, is effected, for example, in accordance with well known methods, by immunizing a mammal, for example a rabbit, with the polypeptide according to the invention or said parts thereof, where appropriate in the presence of, e.g., Freund's adjuvant and/or aluminum hydroxide gels (see, e.g., Diamond, B.A. et al. (1991) The New England Journal of Medicine, 1344). The polyclonal antibodies which have been produced in the animal due to an immunological reaction can then be readily isolated from the blood using well known methods and purified, for example, by way of column chromatography. Preference is given to purifying the antibodies by affinity chromatography, in which, for example, the C-terminal DAN fragment has been coupled to an NHS-activated HiTrap column.

Monoclonal antibodies can, for example, be prepared using the known method of Winter & Milstein (Winter, G. & Milstein, C. (1991) Nature, 349, 293).

The present invention furthermore also relates to a pharmaceutical which comprises a nucleic acid according to the invention or a polypeptide according to the invention and, where appropriate, suitable additives or

adjuvants, and to a process for producing a pharmaceutical for treating cancer, autoimmune diseases, in particular multiple sclerosis or rheumatoid arthritis, Alzheimer's disease, allergies, in particular neurodermatitis, type I allergies or type IV allergies, arthrosis, atherosclerosis, osteoporosis, acute and chronic infectious diseases and/or diabetes, and/or for influencing the metabolism of the cell, in particular in association with immunosuppression, very particularly in association with transplantations, in which pharmaceutical a nucleic acid according to the invention, for example an antisense nucleic acid, or a polypeptide according to the invention is formulated together with pharmaceutically acceptable additives and/or adjuvants.

A pharmaceutical which comprises the nucleic acid according to the invention in naked form or in the form of one of the above-described vectors which are effective in gene therapy, or in the form in which it is complexed with liposomes, is very particularly suitable for gene therapy applications in humans.

Examples of suitable additives and/or adjuvants are a physiological sodium chloride solution, stabilizers, proteinase inhibitors, nuclease inhibitors, etc.

The present invention furthermore also relates to a diagnostic agent which comprises a nucleic acid according to the invention, a polypeptide according to the invention or antibodies according to the invention and, where appropriate, suitable additives and/or adjuvants, and to a process for preparing a diagnostic agent for diagnosing cancer, autoimmune diseases, in particular multiple sclerosis or rheumatoid arthritis, Alzheimer's disease, allergies, in particular neurodermatitis, type I allergies or type IV allergies, arthrosis, atherosclerosis, osteoporosis, acute and chronic infectious diseases and/or diabetes, and/or for analyzing the metabolism of the cell, in particular the immune status, very particularly in association with transplantations, in which diagnostic agent suitable additives and/or adjuvants are added to a nucleic acid according to the invention, a polypeptide according to the invention or antibodies according to the invention.

For example, according to the present invention, the nucleic acid according to the invention can be used to prepare a diagnostic agent which is based on the polymerase chain reaction (PCR diagnostics, e.g. as described in

EP-0200362) or on a Northern blot, as described in more detail in Example 5. These tests are based on the specific hybridization of the nucleic acid according to the invention with the complementary counterstrand, usually of the corresponding mRNA. In this connection, the nucleic acid according to the invention can also be modified, as described, for example, in EP 0063879. Preference is given to labeling a DNA fragment according to the invention with suitable reagents, for example radioactively with α-P^{32}-dATP or nonradioactively with biotin, using well-known methods, and incubating it with isolated RNA which has preferably been bonded to suitable membranes composed, for example, of cellulose or nylon. It is furthermore advantageous to fractionate the isolated RNA according to size, e.g. by means of agarose gel electrophoresis, before hybridization and binding to a membrane. In this way, when the quantity of RNA examined from each tissue sample is the same, it is then possible to determine the quantity of mRNA which has been labeled specifically by the probe.

Another diagnostic agent comprises the polypeptide according to the invention or the immunogenic parts thereof which have been described in more detail above. The polypeptide, or the parts thereof, which are preferably bound to a solid phase, e.g. composed of nitrocellulose or nylon, can, for example, be brought into contact in vitro with the body fluid, e.g. blood, to be investigated in order thereby to be able to react, for example, with autoimmune antibody. The antibody-peptide complex can then, for example, be detected using labeled anti-human IgG or anti-human IgM antibodies. The label is, for example, an enzyme, such as peroxidase, which catalyzes a color reaction. The presence of autoimmune antibodies, and the quantity of these antibodies which is present, can consequently be detected readily and rapidly by way of the color reaction.

Another diagnostic agent comprises the antibodies according to the invention themselves. These antibodies can be used, for example, to readily and rapidly investigate a human tissue sample to determine whether the polypeptide in question is present. In this case, the antibodies according to the invention are, for example, labeled with an enzyme as already described above. The specific antibody-peptide complex can thereby be detected readily and just as rapidly by way of an enzymic color reaction.

The present invention also relates to a test for identifying functional interactors, such as inhibitors or stimulators, comprising a nucleic acid according to the invention, a polypeptide according to the invention or the antibodies according to the invention, and, where appropriate, suitable additives and/or adjuvants.

An example of a suitable test for identifying functional interactors is the so-called two-hybrid system (Fields, S. & Sternglanz, R. (1984) Trends in Genetics, 10, 286). In this test, a cell, for example a yeast cell, is transformed or transfected with one or more expression vectors which express a fusion protein which contains the polypeptide according to the invention and a DNA-binding domain from a known protein, for example from E. coli Gal4 or LexA, and/or express a fusion protein which contains an unknown polypeptide and a transcription-activating domain, for example from Gal4, herpes virus VP16 or B42. In addition, the cell contains a reporter gene, for example the E. coli LacZ gene, "green fluorescence protein" or the yeast amino acid biosynthesis genes His3 or Leu2, which is controlled by regulatory sequences such as the LexA promoter/operator or by a so-called upstream activation sequence (UAS) which is present in the yeast. The unknown polypeptide is, for example, encoded by a DNA fragment which is derived from a gene library, for example from a human gene library. Normally, the above-described expression vectors are used to prepare a cDNA gene library directly in yeast so that the test can be performed immediately thereafter.

For example, the nucleic acid according to the invention is cloned in a yeast expression vector in functional unity onto the nucleic acid encoding the LexA DNA-binding domain, such that the transformed yeast expresses a fusion protein composed of the polypeptide according to the invention and the LexA DNA-binding domain. In another yeast expression vector, cDNA fragments from a cDNA gene library are cloned in functional unity onto the nucleic acid encoding the Gal4 transcription-activating domain, such that the transformed yeast expresses a fusion protein composed of an unknown polypeptide and the Gal4 transcription-activating domain. The yeast which is transformed with the two expression vectors, and which is, for example, Leu2$^-$, additionally contains a nucleic acid which encodes Leu2 and which is controlled by the LexA promoter/operator. If a functional interaction takes place between the polypeptide according to the invention and the unknown polypeptide, the Gal4 transcription-activating domain

then binds, by way of the LexA DNA-binding domain, to the LexA promoter/operator, resulting in the latter being activated and the Leu2 gene being expressed. This then enables the Leu2$^-$ yeast to grow on minimal medium which does not contain any leucine.

When the LacZ reporter gene or the green fluorescence protein reporter gene is used instead of an amino acid biosynthesis gene, activation of transcription can be detected by blue- or green-fluorescent colonies being formed. The blue color or fluorescence color can then be readily quantified in a spectrometer, for example at 585 nM in the case of a blue color.

In this way, it is possible to screen gene expression libraries readily and rapidly for polypeptides which interact with the polypeptide according to the invention. The new polypeptides which have been found can then be isolated and subjected to further characterization.

Another possibility of applying the two-hybrid system is that of using other substances, such as chemical compounds, to influence the interaction between the polypeptide according to the invention and a known or unknown polypeptide. In this way, it is also readily possible to find novel, valuable, chemically synthesizable active compounds which can be employed as therapeutic agents. The present invention is therefore not only directed toward a process for finding polypeptide-like interactors but also extends to a process for finding substances which are able to interact with the above-described protein-protein complex. Within the meaning of the present invention, such peptide-like interactors, and also chemical interactors, are therefore designated functional interactors which are able to have an inhibitory or stimulatory effect.

Another possible application of the polypeptide according to the invention is the poly(A)-specific degradation of nucleic acids, in particular of mRNA. The poly(A)-specific degradation of nucleic acids can be of particular use in research laboratories.

The following figures and examples are intended to clarify the invention without limiting it.

Description of the figures and the most important sequences

SEQ 11 shows the amino acid sequence of the human DAN, containing domains for Exo I (ADFFAIDGEFSGIS), Exo II LVIGHNMLLDVMHTVH) and Exo III (SEQLHEAGYDAYITGLC).

SEQ 12 shows the nucleic acid sequence of the human DAN, including the start codon (pos. 58) and the stop codon (pos. 1977) and the subsequent 3'-UTR.

Fig. 1 shows the elution profile of human DAN on MonoQ (Fig. 3A) and SDS-PAGEs (Figs. 1B and 1C)

Fig. 2 shows a Western blot of recombinant human DAN and native bovine DAN

Fig. 3 shows the deadenylation of mRNA, the accumulation of deadenylated RNA and the influence of PABI

Fig. 4 shows a comparison of known nucleases with the human DAN according to the invention.

Examples

1. Identification of human DAN cDNA clones

Bovine DAN was purified as follows and as described in Körner, G. & Wahle, E. (1997), see above:

All the steps of the purification were carried out at 4°C. Between the purification steps, the samples were frozen in liquid nitrogen and frozen at -80°C. The following basic buffer was used: 50 mM Tris/HCl, 1 mM EDTA, 10% (v/v) glycerol, 1 mM dithiothreitol, 0.4 µg of leupeptin hemisulfate/ml, 0.7 µg of pepstatin/ml, 0.5 mM phenylmethylsulfonyl fluoride, 0.02% (v/v) Nonidet P-40, pH 7.9. Different concentrations of KCl were added to this basic buffer in the various purification steps, as described below.

Fresh calf thymus was obtained from a local slaughterhouse, transported on ice and stored at -80°. One kg was thawed in 2 l of basic buffer containing 50 mM KCl and homogenized in a Waring Blendar homogenizer, initially at low speed, then at medium speed and finally at the

highest speed. The homogenate was centrifuged at 16 000 × g for 1 h and the supernatant was decanted through a wide-mesh gauze (Wahle, E., J. Biol. Chem., 266, 1991).

This calf thymus extract was loaded onto a DEAE-sepharose FF column (column volume 4 l) and eluted from the column at a flow rate of 3 l/h using a salt gradient of from 50 to 600 mM KCl in a volume which was 2.5 times that of the column. Active fractions were eluted from the column at a salt concentration of between 75 and 200 mM KCl. The fractions were collected, combined, treated with ammonium sulfate to give a 30% saturated solution, and stirred on ice for 1.5 h. After a centrifugation (10 800 × g for 30 min, applies to all the centrifugation steps mentioned below), the supernatant was adjusted to 50% saturation with ammonium sulfate and was once again stirred on ice and centrifuged. The sediment was resuspended in 400 ml of basic buffer containing 50 mM KCl and dialyzed against 2 × 4.5 l of basic buffer for 10 h; it was then centrifuged once again. A 1.4 l sepharose-blue column (7 × 36 cm) was equilibrated with basic buffer containing 50 mM KCl and the dialyzed extract was loaded onto the column. The column was washed with 1.5 bed volumes of basic buffer containing 250 mM KCl and eluted with one bed volume of basic buffer containing 1 M KCl (flow rate 2 l/h).

The active fractions from two preparations from in each case 1.2 kg of calf thymus were combined and precipitated with ammonium sulfate (60% saturation). After the centrifugation, the sediment was taken up in 200 ml of basic buffer containing 50 mM KCl, dialyzed for 12 h against 3 × 4 l of the same buffer and loaded, in two portions, onto a heparin-sepharose column (2.5 × 37 cm). The column were washed with 1.5 bed volumes of basic buffer containing 50 mM KCl and then eluted with 10 bed volumes in a gradient of up to 500 mM KCl (flow rate: 145 ml/h). The DAN activity eluted between 80 and 150 mM KCl. The active fractions were combined and dialyzed for 4 h against basic buffer containing 30 mM KCl. The dialysate was centrifuged and chromatographed in two portions on a MonoQ FPLC column (bed volume 8 ml). The column was washed with two bed volumes of basic buffer containing 50 mM KCl and then eluted from the column with a 320 ml gradient containing an increasing concentration of salt (final concentration: 500 mM KCl, flow rate: 2.5 ml/h). The DAN activity eluted at approx. 160 mM KCl. The active fractions (40 ml) were combined and dialyzed for 4 h against 2 l of basic buffer containing 30 mM KCl,

centrifuged and loaded onto a further MonoQ column (1 ml bed volume, flow rate: 0.9 ml/h). The DAN activity bound to the column was eluted with basic buffer containing 500 mM KCl and loaded, in four portions, onto a Superdex HR 10/30 FPLC column (equilibrated with basic buffer containing 300 mM KCl, flow rate: 0.15 ml/h) (Körner, Ch. G. & Wahle, E. (1997), see above).

Different fractions which were eluted from a Superdex column, and which contained the DAN activity, were collected and fractionated through a Phenyl Superose column (Körner, G. & Wahle, E. (1997), see above). The active fractions were collected and dialyzed against buffer (50 mM Tris HCl, pH 7.9, 1 mM EDTA, 10% (v/v) glycerol, 1 mM dithiothreitol, 0.4 μg of leupeptin hemisulfate/ml, 0.7 μg of pepstatin/ml, 0.5 mM phenylmethylsulfonate (PMSF), 0.02% (v/v) Nonidet p-40, 50 mM KCl). After having been centrifuged, the dialysate was loaded onto a 100 μl Smart MonoQ (PC 1.6/5) column. The column was then washed with 200 μl of buffer. The bound proteins were eluted in a gradient containing an increasing concentration of salt (final concentration 500 mM KCl) and having a volume equal to 40 column bed volumes. The fractions which were positive for DAN activity eluted at a KCl concentration of approx. 200 mM KCl. The two fractions having the highest DAN activity were analyzed by SDS polyacrylamide gel electrophoresis. This was then followed by in-situ hydrolysis with the protease Lys-C, fractionation by HPLC and sequencing of the peptides. The following sequences were found:

1. KSFNFYVFPK,
2. KPFNRSSPD(V/K)K,
3. KYAESYWIQTYADYVG and also
4. a mixture composed of: KEQEELNDA and KLFLMRVMD.

It was not possible to determine the N-terminal sequence of the purified bovine DAN.

Subsequently, EST (expressed sequence tags) database searches were carried out using the BLAST program (NCBI). The following clones were identified:

I.M.A.G.E. Consortium clone ID 645295 through peptides 1 and 2, and
I.M.A.G.E. Consortium clone ID 301901 through peptide 3.

The clones which had been found were sequenced completely on an AB1373A sequencing appliance using the ABI Dye Terminator Cycle sequencing kit.

In this connection, it was found that the I.M.A.G.E. consortium clone ID 301901 encodes the 176 C-terminal amino acids of DAN and the entire 3'-UTR, whereas I.M.A.G.E. consortium clone ID 645295 encodes the entire ORF and the 5' UTR and a part of the 3'-UTR, corresponding to a protein composed of 639 amino acids and having a molecular mass of 73.5 kDa (SEQ 11). The 57 nucleotides which are located upstream of the first AUG codon do not contain any stop codon in the reading frame. The sequences (AGAAUGG) surrounding this AUG codon conform to the so-called Kozak rules (Kozak, 1991) which describe preferred starting sequences for the translation start. A 0.7 kB 3'-UTR is present in cDNA clone 645 295 and a poly(A) tail is present at the end of cDNA clone 301901. The poly(A) tail is preceded upstream by the rare polyadenylation signal AUUAAA (SEQ 12).

2. Preparing expression vectors

The plasmid pGMMCS was constructed in the following manner:

The T7 expression vector pGM10 (Martin, 1996), which contained the PABII cDNA sequence (Nemeth, 1996) having an N-terminal Met-Ala-His6 tag, was digested with Xho I and BamH I and the fragment, which contained the 3' moiety of the PAB II, was replaced with a Xho I/BamH I fragment from the multiple cloning site of the pBluescript KS (+/-) plasmid. The resulting plasmid contains a sequence which is regulated by the T7 promoter and begins with Met-Ala His6, followed by the 5' moiety of PAB II and a multiple cloning site. An NDE I cleavage site is located between the His_6 tag and the PAB II sequence and can be used, together with the multiple cloning site, to replace the remaining PAB II sequence with any optional coding sequence.

The plasmid pGMMCS301901 was prepared in the following manner:

A part of the sequence of the human DAN clone containing the 176 C-terminal amino acids was amplified from clone 301901 by PCR using the following conditions:

15 cycles: 45 s at 94°C, 45 s at 54°C and 3.5 min at 72°C. Pfu polymerase and the primers Nde I 301901:

CCATATCCATATGCTTTTCAGTGCCTTTCCTAAC and Xho I 301901: AGTACTCGAGTTACAATGTGTCAGG. Following digestion with Nde I and Xoh I, the resulting cDNA fragments were purified using the Qia-Ex kit (Qiagen GmbH, Hilden) and integrated into the Nde I- and Xho I-digested pGMMCS vector. The sequence was confirmed by sequencing.

The plasmid pGMMCS645295 was prepared as follows:

The coding sequence of the human DAN clone was amplified from clone 645295 by PCR using the following conditions:
20 cycles: 45 s at 94°C, 45 s at 54°C and 3.5 min at 72°C. Pfu polymerase and the primers Nde I 645295:
AGTGTCGCATATGGAGATAATCAGGAGCA and Xho I 645295: AGTACTCAGCGGTTTGCTGCCCTCA. The resulting product was cloned into the vector pCR2.1 using the TA cloning kit (InVitrogen®). Following digestion with Nde I and Xho I, the resulting cDNA fragments were purified using the Qia-Ex kit (Qiagen) and integrated into the Nde I- and Xho I-digested pGMMCS vector. The major part of the PCR-generated sequence was then removed by digesting with Dra III and BstE II and replaced with the corresponding fragment of the original clone. The new sequence was confirmed by sequencing.

3. Preparing antibodies

Antibodies directed against the C-terminal moiety of human DAN were produced as follows:
The plasmid pGMMCS301901 was transformed into BL21 (pLysC). The cells were incubated at 37°C in SOB medium containing 200 mg of carbenicellin ml up to an OD600 of 1.9. 200 mM isopropyl-b-D-thiogalactoside was then added and the mixture was incubated for a further 5 h. The cells were harvested and taken up in buffer A (100 mM NaH_2PO_4, 10 mM Tris, 8M urea, pH 7.9). The cells were lyzed by ultrasonication and the lysate was centrifuged and incubated with 3 ml of Ni-NTA column material at room temperature for 2 h. The column material was packed into a column and washed with 70 ml of buffer A, pH 6.3. It was then washed with 15 ml of buffer B (300 mM NaCl, 10% (v/v) glycerol, 50 mM Tris, 0.01% (v/v) Nonidet-P40, 8M urea, pH 7.9). The protein was refolded on the column by decreasing the urea content with 2 gradients (45 ml/h, gradient 1: buffer B containing from 8 to 4 M urea, at room

temperature, gradient II: containing from 4 to 0 M urea, at 4°C). the protein was eluted from the column with buffer B which contained 500 mM imidazole, dialyzed against buffer B which does not contain any urea and then used for immunizing rabbits.

Anti-DAN antibodies were affinity-purified as follows: the C-terminal DAN fragment was coupled to an NHS-activated HiTrap column (1 ml volume, Pharmacia, Freiburg) in accordance with the manufacturer's instructions. 3 ml of serum were loaded onto the column and eluted in accordance with the manufacturer's instructions. 300 mg of BSA/ml and 0.02% (w/v) NaN_3 were added to the fractions and the buffer was replaced by dialysis.

4. Expressing and purifying human DAN

Human DAN was expressed in E. coli as a fusion protein having an N-terminal His tag (Met-Ala-His_6 tag) in the following manner:

The plasmid pGMMCS645295 was transformed into BL21 (pLysC). The cells were cultured, at 37°C, in 50 ml of LB medium containing 100 mg of carbenicellin/ml and 24 mg of chloramphenicol/ml and then transferred into a 500 ml culture without chloramphenicol and subjected to further incubation at 33°C. After an OD600 at 1.3 had been reached, 100 mM isopropyl-b-D-thiogalactoside was added and the culture was incubated for a further hour. The cells were harvested and taken up in buffer A (50 mM Tris, 300 mM KCl, 0.1 mM MgAc, 1 mM b-mercaptoethanol, 0.4 mg of leupeptin/ml, 0.7 mg of pepstatin/ml, 0.5 mM PMSF, pH 7.9). The cells were lyzed by ultrasonification and the lysate was centrifuged and incubated, at 4°C for 2 h, with 2 ml of Ni^{2+}-NTA column material. The column material was packed into a column and washed with 25 ml of buffer A and then with 20 ml of buffer B (buffer A and 10% (v/v) glycerol, 0.02% (v/v) nonidet-P40, without magnesium, pH 6.3). The protein was then eluted with 5 ml of buffer C (buffer B containing 500 mM imidazole). There then followed a dialysis against buffer D (50 mM Tris, 20 mM KCl, 1 mM EDTA, 5 mM PMSF, 10% (v/v) glycerol, 0.02 (v/v) Nonidet-P40, pH 7.9). The preparation was centrifuged and loaded onto a MonoQ column (1 ml bed volume). The column was then washed with 5 bed volumes of buffer containing 50 mM KCl and eluted with a gradient having a volume equal to 40 bed volumes and a final concentration of 500 mM KCl. The DAN activity eluted in a sharp peak at approx. 190 mM KCl (Fig. 1A).

The purification resulted in a virtually homogeneous preparation of a protein having the expected molecular mass (Figs. 1B and 1C). The protein exhibits a 3'-exoribonuclease activity which is specific for poly(A) (see Example 6). It follows from this that the found cDNA clone encodes a human DAN.

5. Northern blot analysis

RNA was isolated from HeLa cells using a method described by Chomczynski, P. & Sacchi, N. (1987) Anal. Biochem. 162, 156. Poly(A) mRNA was purified using oligo(dT) cellulose in accordance with a method described by Sambrook, J. et al. (1989) Molecular Cloning. A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. The resulting RNA was fractionated on a 1% agarose formaldehyde gel, transferred by capillary blotting to a Hybond N+ membrane (from Amersham Buchler, Brunswick) and fixed on the membrane by incubating at 80°C for 2 h. The sequence encoding human DAN was amplified by PCR, labeled with α-^{32}P-dATP by random priming and used as a probe. The hybridization was carried out as described (Ausubel) and the membrane was washed under high stringency. The Northern blot analysis using HeLa cell poly(A)$^+$ RNA showed one single 3.1 kB fragment which hybridized with the DAN probe. This size agrees well with the size of the cDNA clone.

6. DAN activity test

The assay for DAN activity was carried out as described (Körner, Chr G. & Wahle, E. (1997), see above). SDS polyacrylamide gel electrophoresis (SDS-PAGE) was carried out as described by Laemmli, U.K. (1970) Nature 227, 680.

Like the bovine enzyme, the recombinant DAN is active under two different reaction conditions depending on the neutralization of phosphate charges. In the absence of salt, its activity is completely dependent on the presence of spermidine with the optimum concentration being 1 mM. Under these conditions, the specific activity of the recombinant DAN (79 000 U/mg) is not significantly different from the activity of purified bovine DAN (110 000 U/mg). In the presence of spermidine, the activity of DAN is inhibited by salt. In the absence of spermidine, the activity of the enzyme is

dependent on salt, with the concentration optimum being 150 mM potassium acetate. Bovine DAN behaves in a similar manner under these conditions as well. However, the specific activity of the recombinant protein (960 U/mg) is lower under these conditions than is that of the enzyme purified from calf thymus (8070 U/mg). This difference can be explained either by a post-translational modification or by contaminating proteins being present in the preparation of the bovine enzyme.

When a capped polyadenylated RNA was incubated with the DAN preparations in the presence of salt, the poly(A) tail was degraded and completely deadenylated RNA accumulated transiently (Fig. 3). If the assay was carried out in the presence of PABI, the activity was partially inhibited and step-wise degradation of the poly(A) tail was observed. The length of the predominantly degradation products differed by approx. 30 nucleotides, a phenomenon which was evidently due to the PABI binding. These results concur with the investigations carried out on bovine protein (Körner, Chr. & Wahle, E. (1997), see above). Taken together, these results demonstrate that the recombinant protein is a poly(A)-specific 3'-exoribonuclease.

7. Western blot analyses

Western blot analyses were carried out as follows: the proteins were fractionated by SDS-PAGE and transferred to a nitrocellulose membrane using the semidry method (Kyse-Andersen, 1984). The blots were incubated in TNT buffer (20 mM Tris-HCl, 150 mM NaCl, 0.05% (v/v) Tween 20, pH 7.5) containing 5% (w/v) dried skimmed milk. The same buffer was used for the incubation with antiserum and the washing steps. The blots were incubated with the antibodies at room temperature for 2 - 3 h and then washed. The bound antibodies were detected using a peroxidase-conjugated pig anti-rabbit antibody (DAKO. Glostrup, Denmark) and chemiluminescence staining (SuperSignal kit, Pierce).

The analyses demonstrate that rabbit antibodies which were generated against a C-terminal fragment of DAN are able to precipitate DAN from partially purified fractions. The antibodies recognize both bovine and human DAN in a Western blot (Fig. 4). In this experiment, the recombinant DAN appears to somewhat larger than the enzyme in SDS lysates obtained from HeLa cells, something which can be explained by the

introduction of the tag (1004 Da) into the sequence. If the first AUG in the cDNA sequence had not been the start codon which was used, the natural DAN from HeLa cells would have had to have been at least 1260 Da larger than the tagged recombinant protein. These results, too, confirm that the first AUG sequence in the cDNA clone is used as the start codon in vivo.